

**MAT²**

MATerials MATemàtics

Volum 2020, treball no. 3, 35 pp. ISSN: 1887-1097

Publicació electrònica de divulgació del Departament de Matemàtiques
de la Universitat Autònoma de Barcelonawww.mat.uab.cat/web/matmat

Una passejada aleatòria pel conjunt de Cantor

Frederic Utzet

1 Introducció

La funció de Cantor, que es construeix a partir del conjunt de Cantor, és una funció $F : [0, 1] \rightarrow [0, 1]$ contínua, no decreixent, amb $F(0) = 0$ i $F(1) = 1$, però amb derivada zero en quasi tots els punts. Aquestes propietats semblen incompatibles: ¿com pot ser que vagi del $(0,0)$ al $(1,1)$ amb continuïtat essent localment constant en quasi tots els punts? ¿On es produeix el creixement? És difícil d'imaginar i impossible de dibuixar, i també és coneguda com *l'escala del diable* (Mandelbrot [18]). Vegeu-ne un dibuix aproximat a la Figura 1.

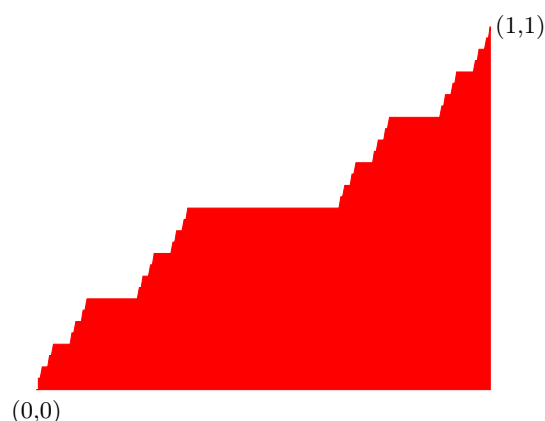


Figura 1: Dibuix aproximat de l'escala del diable

En aquest article es posa el focus en la **distribució de probabilitat de Cantor**, la qual es defineix mitjançant la funció de Cantor. He intentat recollir de manera coherent i autocontinguda les seves propietats més importants, les quals estan disperses en els llibres com problemes o notes a peu

de pàgina, sovint sense cap demostració. Aquesta distribució de probabilitat té trets molt poc habituals, però és sorprenent que, malgrat que no és discreta ni té densitat, es puguin estudiar de forma senzilla moltes de les seves propietats; per exemple, trobar-ne l'esperança, o, en general, els moments de qualsevol ordre, o calcular-ne la funció característica. A més, és especialment interessant que es pot obtenir com la llei del límit d'una passejada aleatòria sobre el conjunt de Cantor, i quan s'adopta aquesta perspectiva les seves propietats es tornen més clares i intuïtives.

Voldria assenyalar que la funció i la distribució de Cantor són molt conegudes com a curiositats o com fonts de contraexemples. Però també, i més important, han estimulat la imaginació de molts matemàtics; entre ells i en un lloc destacat, cal esmentar el genial Paul Erdős, amb els articles [9] i [10] que van tenir una influència profunda en el desenvolupament de diverses branques de les matemàtiques; vegeu la revisió històrica de Peres *et al.* [21].

L'article està dividit en quatre parts. La primera està dedicada al conjunt de Cantor (seccions 2 a 4). La segona a la funció de Cantor (secció 5 i 6). La tercera a la distribució de Cantor i les seves propietats deduïdes a partir de la seva funció de distribució (seccions 7 a 10). A la quarta part s'introdueixen breument les passejades aleatòries i es recupera la distribució de Cantor com a distribució del límit d'una passejada aleatòria sobre el conjunt de Cantor i es presenten demostracions alternatives de les seves propietats (seccions 11 a 13).

2 El conjunt de Cantor

El conjunt de Cantor s'obté eliminant sistemàticament fragments de l'interval $[0, 1]$; específicament, posem

$$C_0 = [0, 1].$$

Dividim C_0 en tres intervals de longitud $1/3$:

$$C_0 = [0, \frac{1}{3}] \cup (\frac{1}{3}, \frac{2}{3}) \cup [\frac{2}{3}, 1].$$

Anomenem C_1 al conjunt que queda després de treure'n la part central:

$$C_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1].$$

Ara dividim $[0, \frac{1}{3}]$ en tres intervals de mida $1/9$ i en traiem la part central, i fem el mateix amb $[\frac{2}{3}, 1]$. El resultat l'anomenem C_2 :

$$C_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1].$$

D'aquesta manera anem construint una successió decreixent de conjunts (vegeu la Figura 2)

$$C_0 \supset C_1 \supset C_2 \dots$$

S'anomena **conjunt de Cantor** al límit d'aquesta successió de conjunts i el designarem per C ; més formalment, el conjunt de Cantor és la intersecció de tots aquests conjunts:

$$C = \bigcap_{n=0}^{\infty} C_n.$$

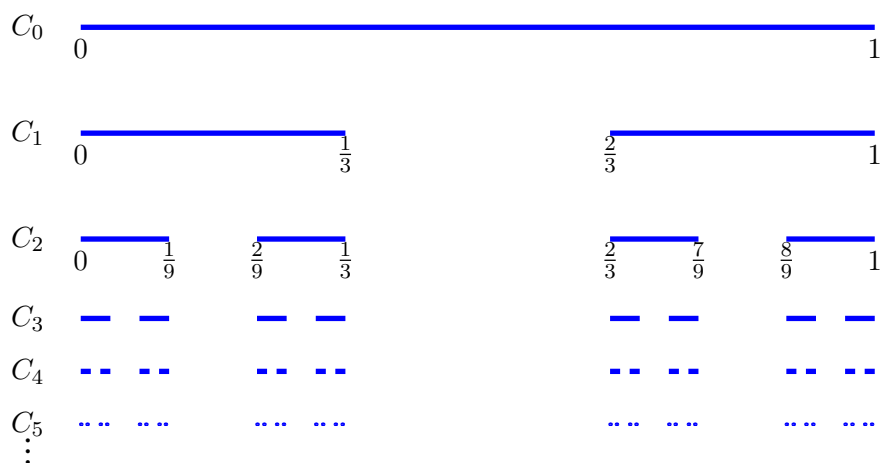


Figura 2: Construcció del conjunt ternari de Cantor

¿Queda alguna cosa, al conjunt C ? Sembla que només hi ha una col·lecció de forats. Mandelbrot [18] deia que queda *la pols de Cantor*¹ i l'expressió està certament ben trobada i té gràcia. Però com veurem, aquesta *pols* té moltes partícules. A la Figura 2 observem el caràcter *fractal* o *autosimilar* del conjunt C : si ampliem tres cops la part de C continguda a $[0, \frac{1}{3}]$ obtenim tot el conjunt C , o si augmentem 9 vegades la part de C a l'interval $[\frac{6}{9}, \frac{7}{9}]$ tornem a obtenir tot C .

Segons explica Fléron [13] aquesta mena de conjunts obtinguts foradant iterativament un interval de \mathbb{R} , que actualment es coneixen com a conjunts de Cantor, va ser introduïda per J. S. Smith en un article l'any 1875. Cantor, l'any 1883, va ser el primer en construir el conjunt que hem presentat, i que més pròpiament s'anomena *conjunt ternari de Cantor* o també *conjunt del terç intermedi* o *conjunt discontinu de Cantor*. Cantor va definir aquest conjunt en el context de la topologia dels nombres reals i a partir d'ell va construir una funció *patològica* que va utilitzar de contraexemple en l'extensió que en aquell temps s'estava fent del Teorema fonamental del càlcul a funcions discontinues (Fléron [13]).

Per completar els detalls històrics, Pesin and Climenhaga [22, p. 17] comenten que Cantor creia que el conjunt ternari tindria un cardinal que estaria estrictament entre el dels nombres naturals (\aleph_0) i el dels nombres reals

¹El terme *Pols de Cantor* també s'utilitza per designar una construcció bidimensional anàloga al conjunt de Cantor, vegeu, per exemple, Falconer [11, p. XX]

(\aleph_1). Però el mateix Cantor va demostrar que el conjunt ternari tenia igual cardinal que \mathbb{R} , i llavors va conjeturar –hipòtesi del continu– la inexistència d'aquest cardinal intermedi, i va intentar obsessiva, però infructuosament, provar-ho durant anys; demostrar aquesta conjectura és el primer problema de la llista de Hilbert de 1900. La seva indemostrabilitat va ser establerta per Paul J. Cohen l'any 1963 (vegeu Cohen [6]) afinant eines inventades per Gödel; Cohen va rebre la medalla Fields el 1966 per aquest resultat.

2.1 Notació dels intervals

Utilitzarem la següent notació. Fixat un nivell n , escriurem

$$I_k^n, \quad k = 1, \dots, 2^n,$$

els intervals tancats que formen C_n escrits en l'ordre que apareixen d'esquerra a dreta en un dibuix, vegeu la Figura 3. Designarem per

$$E_k^n, \quad k = 1, \dots, 2^n - 1,$$

els intervals que hem exclòs a tots els nivells fins al n (inclòs). Llavors (vegeu la Figura 3),

$$[0, 1] = I_1^n \cup E_1^n \cup I_2^n \cup \dots \cup I_{2^n-1}^n \cup E_{2^n-1}^n \cup I_{2^n}^n.$$

És important observar que I_k^n és de la forma $\left[\frac{j}{3^n}, \frac{j+1}{3^n}\right]$ per algun índex $j \in \{0, \dots, 2^n - 1\}$, però que

$$I_k^n \neq \left[\frac{k}{3^n}, \frac{k+1}{3^n}\right].$$

Noteu també que fixat $n \geq 1$, tots els intervals $I_1^1, \dots, I_{2^n}^n$ són de mida 3^{-n} , però que $E_1^n, \dots, E_{2^n-1}^n$, són de diferents mides ja que procedeixen de totes les etapes $1, \dots, n$.

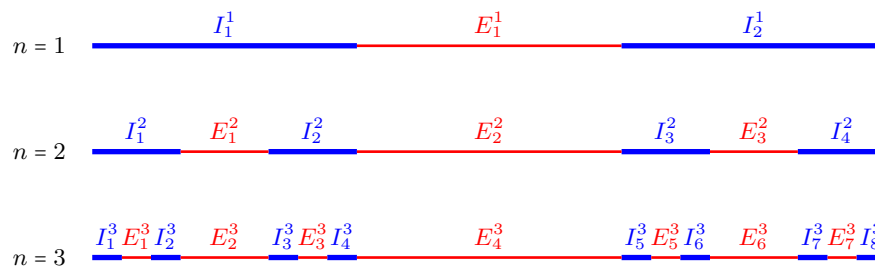


Figura 3: Els intervals I_k^n i E_k^n

3 Representacions binàries i ternàries dels nombres decimals

Per tal d'estudiar les propietats del conjunt de Cantor és convenient utilitzar la representació dels nombres de $[0, 1]$ en base 3. Començarem repassant breument el sistema binari.

3.1 El sistema binari

Recordem que un nombre natural b es pot representar en base 2 de la forma

$$b = b_n \times 2^n + \dots + b_1 \times 2 + b_0, \quad \text{amb } b_0, \dots, b_n \in \{0, 1\},$$

i s'escriu

$$b = (b_n b_{n-1} \dots b_1 b_0)_2.$$

Per exemple,

$$9 = 1 \times 2^3 + 0 \times 2^2 + 0 \times 2 + 1 = (1001)_2.$$

Però també es poden representar en base 2 els nombres decimals: qualsevol nombre $a \in [0, 1]$ pot escriure's de la forma

$$a = a_1 \times \frac{1}{2} + a_2 \times \frac{1}{2^2} + \dots = \sum_{n=1}^{\infty} \frac{a_n}{2^n}, \quad \text{amb } a_1, a_2, \dots \in \{0, 1\},$$

i posarem

$$a = (0.a_1 a_2 \dots)_2.$$

(Per motius de claredat tipogràfica utilitzarem el punt -que en aquest cas s'anomena punt binari- en lloc de la coma normativa de la notació decimal.)

Per exemple,

$$0,6875 = \frac{11}{16} = \frac{1}{2} + \frac{1}{2^3} + \frac{1}{2^4} = (0.1011)_2.$$

3.1.1 Nombres amb dues expressions binàries

A l'igual que hi ha nombres amb dues expressions decimals, per exemple, $0,5 = 0,49$, hi ha nombres de l'interval $[0, 1]$ amb dues expressions binàries. Per exemple, tenint en compte la suma de la progressió geomètrica de raó $1/2$,

$$\frac{1}{2} + \frac{1}{2^2} + \dots = \sum_{n=1}^{\infty} \frac{1}{2^n} = 1, \quad (1)$$

es comprova que

$$0,5 = (0.1)_2 = (0.0\hat{1})_2.$$

L'expressió $(0.1)_2$ s'anomena *expressió finita*, mentre que $(0.0\hat{1})_2$ s'anomena *expressió infinita*. Tots els nombres de la forma $k/2^n$, $k = 1, \dots, 2^n - 1$, $n \geq 1$, i només aquests nombres, tenen dos desenvolupaments. Aquest fet normalment no comporta cap problema, però cal tenir-lo present quan es vol identificar de forma única cada punt de l'interval $[0, 1]$ amb una expressió binària.

3.2 El sistema ternari

De manera completament anàloga, tot nombre $a \in [0, 1]$ es pot representar en base 3: tenim que

$$a = \sum_{n=1}^{\infty} \frac{a_n}{3^n}, \quad \text{amb } a_n \in \{0, 1, 2\},$$

que abreujadament s'escriu

$$a = (0.a_1a_2\cdots)_3,$$

i que s'anomena **expressió ternària** del nombre a (també utilitzarem el *punt ternari* en lloc de la coma).

Els nombres de la forma $k/3^n$, amb $1 \leq k \leq 3^n - 1$ (i només aquests nombres) tenen dues expressions, una només amb zeros a partir de cert lloc (expressió finita), i l'altra només amb dosos a partir de cert lloc (expressió infinita). Per exemple,

$$\frac{1}{3} = (0.1)_3 = (0.0\widehat{2})_3 \quad \text{i} \quad \frac{2}{3} = (0.2)_3 = (0.1\widehat{2})_3.$$

Utilitzarem sovint la suma d'una progressió geomètrica de raó $1/3$. Per posterior referència, recordem que

$$\sum_{j=1}^{\infty} \frac{1}{3^j} = \frac{\frac{1}{3}}{1 - \frac{1}{3}} = \frac{1}{2}. \quad (2)$$

D'aquí (o directament per la fórmula de la suma d'una progressió geomètrica),

$$\sum_{j=n+1}^{\infty} \frac{1}{3^j} = \frac{1}{3^n} \sum_{j=1}^{\infty} \frac{1}{3^j} = \frac{1}{2} \frac{1}{3^n}. \quad (3)$$

4 Algunes propietats del conjunt de Cantor

Ens limitarem a les propietats que utilitzarem més endavant. Per les propietats topològiques del conjunt de Cantor vegeu Agudé [1, pp. 43–44 i 51–53] o per les relacions amb els sistemes dinàmics Haro [15].

4.1 Representació ternària dels elements del conjunt de Cantor

La notació ternària s'adapta molt bé a l'estructura del conjunt de Cantor i tot és molt més clar quan s'escriuen els nombres en base 3. Ens proposem

argumentar que el conjunt de Cantor està format pels nombres que tenen una expressió ternària formada per zeros i dosos:

$$a \in C \iff a = (0.a_1a_2, \dots)_3, \quad \text{amb } a_n \in \{0, 2\}, \quad \forall n \geq 1. \quad (4)$$

En efecte, els nombres de l'interval $[0, 1/3]$ tindran 0 com a primera xifra de la seva expressió ternària, i els de l'interval $[2/3, 1]$ tindran 2 com a primera xifra. Així, amb les notacions de la secció 2, els nombres del conjunt C_1 tindran com a primera xifra (després del punt decimal) de la seva expressió ternària 0 o 2. Anàlogament, els de C_2 , després del punt, tindran les dues primeres xifres de la forma 00, 02, 20 o 22. I així successivament. Deixarem al lector la formalització d'aquest raonament.

Exemple. Mirant la Figura 2 sembla que C només està format pels extrems dels intervals dels conjunts C_n ; aquests extrems són de la forma $k/3^n$ per certs valors de k ; però el conjunt d'aquests nombres és numerable i de seguida veurem que C no ho és, de manera que té molts altres punts; per exemple, $1/4 \in C$, ja que la seva representació en base 3 és $(0.\overline{02})_3$. Però $1/4$ no és l'extrem de cap interval de C_n ja que és clar que no es poden trobar $k, n \in \mathbb{N}$ tals que

$$\frac{1}{4} = \frac{k}{3^n}.$$

4.2 El conjunt de Cantor té mesura de Lebesgue zero

La mesura de Lebesgue a \mathbb{R} , que designarem per \mathcal{L} , és l'extensió de la longitud d'un segment a conjunts de \mathbb{R} més generals, i es caracteritza pel fet que un interval (a, b) mesura $b - a$:

$$\mathcal{L}(a, b) = b - a, \quad a \leq b.$$

Llavors, és senzill calcular la mesura de Lebesgue de C_n , ja que està format per 2^n intervals disjunts cadascun de longitud $1/3^n$; així, d'acord amb les notacions de la secció 2.1,

$$C_n = \bigcup_{k=1}^{2^n} I_k^n,$$

i llavors

$$\mathcal{L}(C_n) = \sum_{k=1}^{2^n} \mathcal{L}(I_k^n) = 2^n \cdot \frac{1}{3^n}.$$

De la definició de C i de les propietats de les mesures,

$$\mathcal{L}(C) = \lim_n \mathcal{L}(C_n) = \lim_n \left(\frac{2}{3}\right)^n = 0.$$

4.3 El conjunt de Cantor té el mateix cardinal que $[0,1]$

De la caracterització de C com els elements que en base 3 només tenen zeros i dosos es dedueix que C i $[0,1]$ tenen el mateix cardinal. En efecte, d'una banda, atès que $C \subset [0,1]$ deduïm que el cardinal de C és menor que el de $[0,1]$: $\text{Card}(C) \leq \text{Card}([0,1])$.

Per veure l'altra desigualtat, considerem l'aplicació $H : C \rightarrow [0,1]$ definida per

$$H((0.x_1x_2\dots)_3) = (0.y_1y_2\dots)_2, \quad \text{on } y_n = x_n/2, \quad \forall n \geq 1.$$

Vegeu la Figura 4. Clarament, la funció H és exhaustiva, d'on $\text{Card}(C) \geq \text{Card}([0,1])$.

La funció H es pot escriure de manera més compacta com

$$H((0.x_1x_2\dots)_3) = \sum_{n=1}^{\infty} \frac{x_n}{2^{n+1}}. \quad (5)$$

Cal remarcar que H no és injectiva. Per exemple,

$$H(1/3) = H((0.0\widehat{2})_3) = (0.0\widehat{1})_2 = 1/2.$$

Però també

$$H(2/3) = H((0.2)_3) = (0.1)_2 = 1/2.$$

La funció H és contínua; les aparents discontinuïtats de la Figura 4 són degudes (a part que el gràfic de H no es pot dibuixar) al fet que la funció només està definida a C ; per exemple, l'interval $(\frac{1}{3}, \frac{2}{3})$ no està inclòs en C . Estem acostumats a treballar amb funcions definides en un interval on la continuïtat es «veu» gràficament.

Demostració de la continuïtat de H

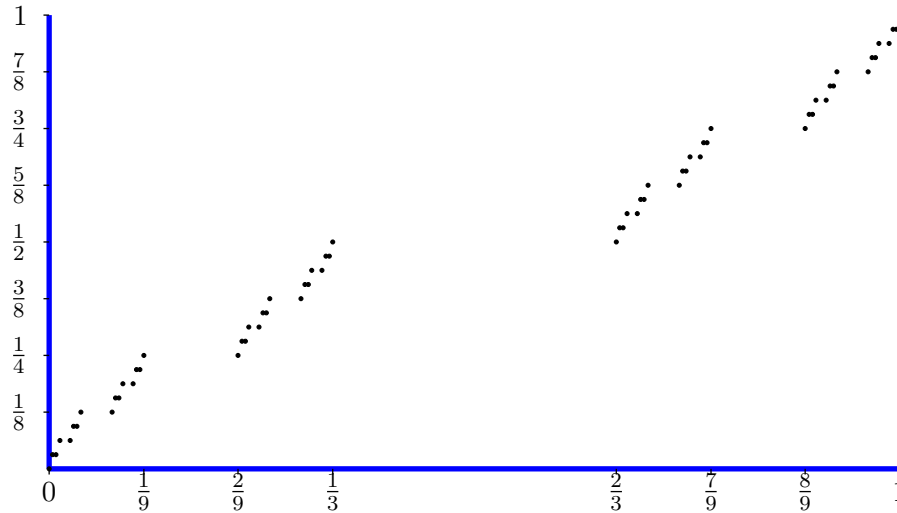
La continuïtat de H es demostra a partir del fet que la proximitat entre dos punts $x = (0.x_1x_2\dots)_3$ i $y = (0.y_1y_2\dots)_3$ es tradueix en què per a cert n , tenim que $x_j = y_j$, $j = 1, \dots, n$. De la definició de H es segueix que llavors si x i y estan prou propers, també ho estaran les seves imatges.

Formalment, fixem $\varepsilon > 0$ i prenem n_0 tal que $2^{-n_0} < \varepsilon$. Siguin $x = (0.x_1x_2\dots)_3$ i $y = (0.y_1y_2\dots)_3$ dos punts de C tals que

$$|x - y| < \frac{1}{3^{n_0}}. \quad (6)$$

Llavors

$$x_i = y_i, \quad i = 1, \dots, n_0,$$

Figura 4: Gràfic aproximat de la funció $H : C \rightarrow [0, 1]$.

ja que, en cas contrari, si per a $j < n_0$ tinguéssim $x_1 = y_1, \dots, x_{j-1} = y_{j-1}$ i $x_j \neq y_j$, per exemple, $x_j = 0$ i $y_j = 2$, llavors

$$\begin{aligned} |x - y| &= y - x \geq (0.y_1 \dots y_{j-1}2)_3 - (0.x_1 \dots x_{j-1}1)_3 \\ &= (0.\underbrace{0 \dots 0}_{j-1}1)_3 = \frac{1}{3^j} > \frac{1}{3^{n_0}}, \end{aligned}$$

que contradiria (6).

Aleshores,

$$|H(x) - H(y)| \leq \sum_{j=n_0+1}^{\infty} \frac{1}{2^j} = \frac{1}{2^{n_0}} < \varepsilon.$$

5 La funció de Cantor

La funció de Cantor $F : [0, 1] \rightarrow [0, 1]$ és l'extensió de la funció $H : C \rightarrow [0, 1]$ a tot $[0, 1]$. És molt fàcil construir-la gràficament, ja que H assigna el mateix valor a ambdós extrems dels intervals E_k^n que hem eliminat a la construcció de C ; llavors, senzillament, tracem una línia horitzontal sobre aquest interval, compareu la Figura 4 amb la Figura 5. La construcció de H i la seva extensió F van ser fetes per Cantor [3]. Però és un repte escriure explícitament la funció F ; la fórmula que donem es troba, per exemple, a Halmos [14, p. 83]. Finalment, veurem també una construcció de F com a límit d'una successió de funcions contínues.

5.1 La funció de Cantor com extensió de H

Tal com hem dit, la funció de Cantor $F : [0, 1] \rightarrow [0, 1]$ es construeix estenent H a tot l'interval $[0, 1]$ de la següent manera: si $E_k^n = (a, b)$ aleshores per a qualsevol $x \in E_k^n$, definim

$$F(x) = H(a) = H(b).$$

Concretament, la funció de Cantor val $1/2$ al primer interval que hem suprimit al passar de C_0 a C_1 , és a dir, val $1/2$ a $E_1^1 = (1/3, 2/3)$. Val $1/4$ al segon interval escapçat, $E_1^2 = (1/9, 2/9)$, i $3/8$ a $E_3^2 = (7/9, 8/9)$; $1/8$ a $E_1^3 = (1/27, 2/27)$ i així successivament. La Figura 1 que hem vist a la introducció reproduïx (aproximadament) un dibuix de Mandelbrot [18] que afegeix dramatisme al gràfic de la funció F .

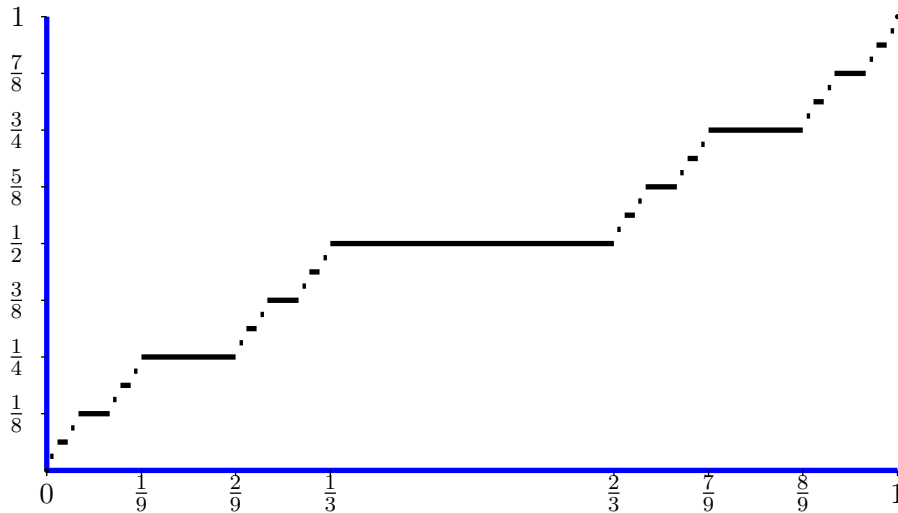


Figura 5: Gràfic aproximatiu de la funció de Cantor $F : [0, 1] \rightarrow [0, 1]$

5.2 Definició explícita de la funció de Cantor

Quan $x \in C$, com que F és l'extensió de H , tenim $F(x) = H(x)$, on H està definida a (5); és a dir, per $x = (0.x_1x_2\cdots)_3 \in C$,

$$F((0.x_1x_2\cdots)_3) = \sum_{n=1}^{\infty} \frac{x_n}{2^{n+1}}. \quad (7)$$

Quan $x \notin C$, aleshores x serà d'algun conjunt E_k^n , ja que

$$[0, 1] \setminus C = \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{2^n-1} E_k^n.$$

Llavors, si tenim que $x \in E_k^n$, i posem que aquest interval és $E_k^n = (a, b)$, amb $a, b \in C$, tal com hem fet a l'apartat anterior, es defineix

$$F(x) = H(a) = H(b).$$

Aleshores la dificultat està en identificar els extrems de l'interval E_k^n (de fet, amb l'extrem inferior n'hi ha prou). Això es fa de la següent manera: si $x = (0.x_1x_2\dots)_3 \notin C$ contindrà algun 1 en la seva expressió ternària; designem per $\ell(x)$ el lloc que ocupa el primer 1:

$$\ell(x) = \text{primer índex } n \text{ tal que } x_n = 1.$$

Així, (escric ℓ en lloc de $\ell(x)$)

$$x = (0.x_1 \dots x_{\ell-1} 1 x_{\ell+1} \dots)_3, \text{ amb } x_1, \dots, x_{\ell-1} \in \{0, 2\}.$$

El lector ha de convèncer-se que l'extrem inferior a de l'interval al que ens hem referit abans és

$$a = (0.x_1 \dots x_{\ell-1} 1)_3.$$

Però aquest nombre també té una expressió només amb zero i dosos (recordeu que $a \in C$), que serà

$$a = (0.x_1 \dots x_{\ell-1} 0\widehat{2})_3.$$

Llavors, d'acord amb (7),

$$F(x) = F(a) = H(a) = \sum_{n=1}^{\ell-1} \frac{x_n}{2^{n+1}} + \sum_{m=\ell+1}^{\infty} \frac{1}{2^m} = \sum_{n=1}^{\ell-1} \frac{x_n}{2^{n+1}} + \frac{1}{2^\ell}. \quad (8)$$

Ambdues expressions (7) i (8) poden unificar-se definint $\ell(x) = \infty$ quan $x \in C$. Llavors per a $x = (0.x_1x_2\dots)_3 \in [0, 1]$

$$F(x) = \sum_{n=1}^{\ell(x)-1} \frac{x_n}{2^{n+1}} + \frac{1}{2^{\ell(x)}}. \quad (9)$$

Cal notar que, quan un nombre x té dues expressions ternàries, aquesta fórmula els hi dona el mateix valor (observeu que $\ell(x)$ sí que depèn de l'expressió ternària, però $F(x)$ no).

5.3 La funció de Cantor com a límit d'una successió de funcions senzilles

No utilitzarem aquesta construcció, però és interessant comentar-la breument. Es fa mitjançant una successió de funcions contínues que aproximem

a F . Amb les notacions que hem introduït a la subsecció 2.1, definim

$$F_n(x) = \begin{cases} 0, & \text{si } x = 0, \\ \frac{k}{2^n}, & \text{si } x \in E_k^n, \quad k = 1, \dots, 2^n - 1, \\ 1, & \text{si } x = 1. \end{cases} \quad (10)$$

I interpolem linealment entre els punts indicats. Vegeu la Figura 6.

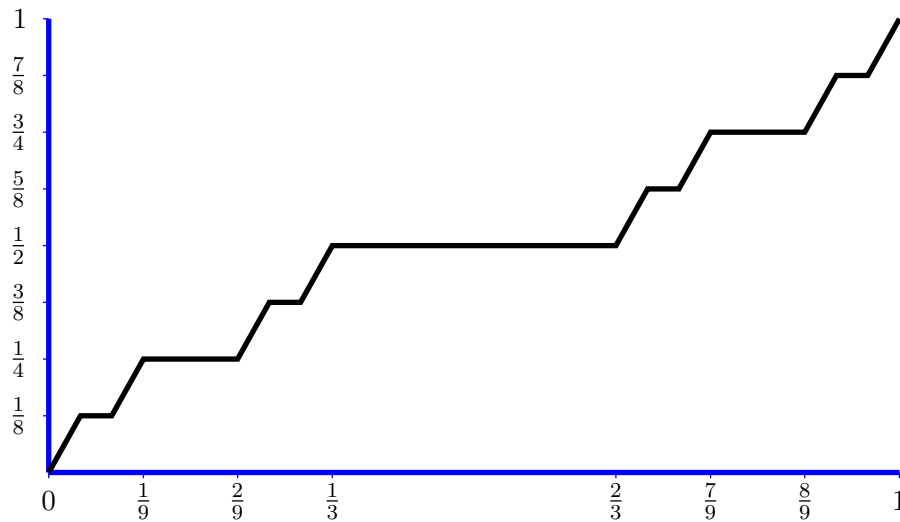


Figura 6: Gràfic de la funció F_3

Gràficament sembla clar que

$$\lim_n F_n(x) = F(x), \text{ per tot } x \in [0, 1],$$

però la demostració és força laboriosa i no la farem.

6 Propietats de la funció de Cantor

En aquesta secció veurem les propietats que utilitzarem directament en l'estudi de la distribució de Cantor. La funció de Cantor té moltes altres propietats, vegeu, per exemple, Dovgosheya *et al.* [8].

6.1 $F(0) = 0$ i $F(1) = 1$

6.2 F és no decreixent

Si $x < y$, llavors $F(x) \leq F(y)$. En efecte, siguin $x = (0, x_1 x_2 \dots)_3$ i $y = (0, y_1 y_2 \dots)_3$. Llavors $x < y$ vol dir que si tenim $x_i = y_i$ per a $i = 1, \dots, n$, i

$x_{n+1} \neq y_{n+1}$, llavors $x_{n+1} < y_{n+1}$, d'on analitzant els diferents casos es dedueix que $F(x) \leq F(y)$.

6.3 F és contínua

En efecte, si $x \in [0, 1] \setminus C$, llavors, tal com ja hem utilitzat, $x \in E_k^n$ per alguns $n \geq 1$ i $k \in \{1, \dots, 2^n - 1\}$. Però F és constant en aquest interval, i per tant contínua en x .

Si $x \in C$, aleshores la continuïtat en x es dedueix fàcilment de la continuïtat de H en x i del caràcter no decreixent de F .

6.4 F és derivable en quasi tots els punts i (en aquests punts) $F'(x) = 0$

Si $x \in [0, 1] \setminus C$, llavors F és derivable en x i $F'(x) = 0$. En efecte, tal com hem vist a la demostració de la propietat anterior, F és constant en un interval de x (es diu que F és localment constant en x), d'on $F'(x) = 0$. Atès que la mesura de Lebesgue de C és 0, tenim que F és derivable i té derivada 0 en quasi tot punt (respecte la mesura de Lebesgue) de $[0, 1]$.

Com a conseqüència d'això, F no es pot obtenir com la integral de la seva derivada:

$$F(x) \neq \int_0^x F'(t) dt = 0.$$

Tècnicament es diu que F és contínua però no absolutament contínua.

6.5 El gràfic de la funció F és simètric respecte el punt $(\frac{1}{2}, \frac{1}{2})$

En efecte, dos punts del pla (x, y) i (x', y') són simètrics respecte el punt $(\frac{1}{2}, \frac{1}{2})$ si i només si

$$\begin{aligned} x' &= 1 - x, \\ y' &= 1 - y. \end{aligned}$$

Aleshores, si prenem un punt $(x, F(x))$ del gràfic de F , el seu simètric respecte $(\frac{1}{2}, \frac{1}{2})$ serà $(1 - x, 1 - F(x))$ (vegeu la Figura 7); llavors, per demostrar que el gràfic és simètric hem de comprovar que

$$F(1 - x) = 1 - F(x),$$

o, equivalentment,

$$F(x) + F(1 - x) = 1. \quad (11)$$

Per demostrar aquesta igualtat, considerem $x = 0.x_1x_2 \dots \in [0, 1]$ (en base 3, no escric els subíndexs). D'acord amb les notacions de la secció 5, sigui $\ell(x)$ el primer índex on apareix la xifra 1, i posem $\ell(x) = \infty$ si $x \in C$. Donat que $1 = 0.\bar{2}$, tindrem que $1 - x = 0.(2 - x_1)(2 - x_2) \dots$, i $\ell(1 - x) = \ell(x)$. Llavors, aplicant la definició de F donada per (9), és comprova (11).

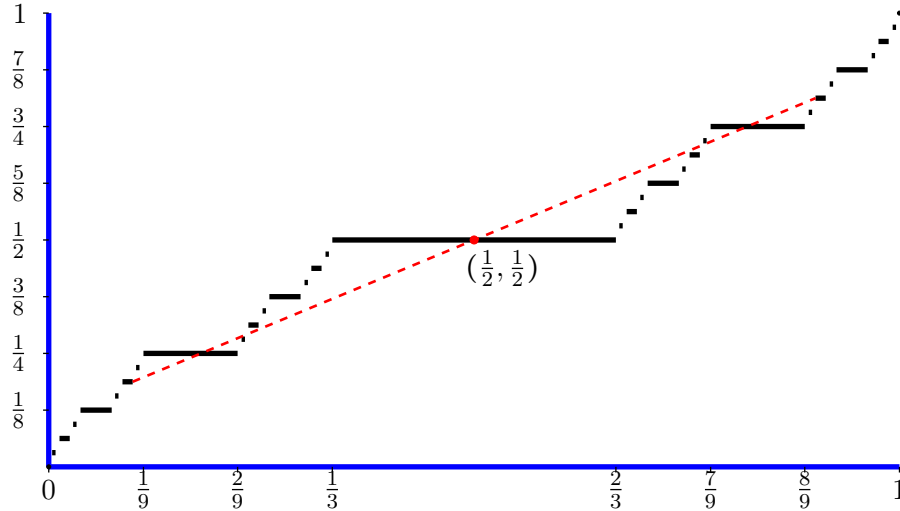


Figura 7: El gràfic de la funció F és simètric respecte el punt $(\frac{1}{2}, \frac{1}{2})$.

6.6 Una caracterització de la funció de Cantor

La igualtats següents són directes de la definició de F (vegeu Chung [5, p. 15], però són interessants perquè en certa manera caracteritzen F (vegeu l'enunciat precís a Dovgosheya *et al.* [8, §4,7]) i juguen un paper important en les demostracions de les propietats de la distribució de Cantor:

$$F\left(\frac{x}{3}\right) = \frac{1}{2} F(x) \quad \text{i} \quad F\left(\frac{x}{3} + \frac{2}{3}\right) = \frac{1}{2} + \frac{1}{2} F(x). \quad (12)$$

7 La distribució de Cantor

En el que segueix s'utilitzen els conceptes i les notacions habituals de la teoria de la probabilitat; per una presentació general d'aquesta teoria el lector pot consultar Sanz [24].

7.1 Variables aleatòries i funcions de distribució

Recordem que a una variable aleatòria Y li correspon la seva funció de distribució $G: \mathbb{R} \rightarrow [0, 1]$ definida per

$$G(t) = \mathbf{P}(Y \leq t) = \mathbf{P}(Y \in (-\infty, t]), \quad t \in \mathbb{R}.$$

La funció G determina totes les probabilitats relatives a Y ; per exemple, per qualsevol $-\infty \leq a < b \leq \infty$,

$$\mathbf{P}(a < Y \leq b) = \mathbf{P}(Y \in (a, b]) = G(b) - G(a). \quad (13)$$

O bé,

$$\mathbf{P}(Y = a) = \mathbf{P}(Y \in \{a\}) = G(a) - G(a^-), \quad (14)$$

on $G(a^-)$ és el límit per l'esquerra de G en el punt a .

La funció G té les següents propietats:

1. És no decreixent: si $x < y$ aleshores $G(x) \leq G(y)$.
2. És contínua per la dreta i té límit per l'esquerra en tots els punts.
3. $\lim_{x \rightarrow -\infty} G(x) = 0$ i $\lim_{x \rightarrow +\infty} G(x) = 1$.

Recíprocament, donada una funció G que compleixi 1, 2 i 3, aleshores existeix una variable aleatòria Y (en algun espai de probabilitat) tal que G és la seva funció de distribució.

7.2 Definició de la distribució de Cantor

La funció de Cantor $F : [0, 1] \rightarrow [0, 1]$ es pot estendre a tot \mathbb{R} definint $F(x) = 0$, si $x < 0$ i $F(x) = 1$, si $x > 1$; designarem aquesta extensió per la mateixa lletra F . Aleshores, F compleix 1, 2 i 3 de l'apartat anterior, i per tant és una funció de distribució. Designem per X una variable aleatòria que tingui funció de distribució F ; es diu que X té (o segueix) la **distribució de Cantor**. Del fet que F sigui contínua, es dedueix que per a qualsevol x ,

$$F(x^-) = F(x).$$

Llavors, per (14),

$$\mathbf{P}(X = x) = 0. \quad (15)$$

D'on

$$\begin{aligned} \mathbf{P}(a < X < b) &= P(a \leq X \leq b) = \mathbf{P}(a < X \leq b) = \mathbf{P}(a \leq X < b) \\ &= F(b) - F(a). \end{aligned} \quad (16)$$

A continuació, veurem com es trameten les propietats de F a X .

8 Propietats de la distribució de Cantor

En aquesta secció, X sempre és una variable aleatòria amb distribució de Cantor.

8.1 La distribució de Cantor està concentrada en el conjunt de Cantor

Volem veure que $\mathbf{P}(X \in C) = 1$, o, equivalentment, que $\mathbf{P}(X \in C^c) = 0$. Pel lector familiaritzat amb les funcions de distribució aquesta propietat és molt evident, ja que F és constant en C^c (vegeu la Figura 5). Concretament,

$$C^c = (\infty, 0) \cup (1, \infty) \cup \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{2^n-1} E_k^n,$$

i per qualsevol d'aquests intervals, posem (a, b) , tenim

$$\mathbf{P}(X \in (a, b)) = F(b) - F(a) = 0.$$

D'on es dedueix,

$$\mathbf{P}(X \in C^c) = 0.$$

Atès que el conjunt de Cantor de mesura de Lebesgue zero (vegeu la secció 4.2), és diu que la distribució de Cantor és **singular** (respecte la mesura de Lebesgue).

8.2 Equiprobabilitat en cada nivell

El conjunt C_n de la construcció que hem fet a la secció 2 està format per 2^n intervals disjunts I_k^n , $k = 1, \dots, 2^n$; anem a provar que tots aquests intervals tenen la mateixa probabilitat. Concretament, per qualsevol $n \geq 1$,

$$\mathbf{P}(X \in I_k^n) = \frac{1}{2^n}, \quad k = 1, \dots, 2^n. \quad (17)$$

En efecte, cada interval I_k^n té de longitud 3^{-n} . Notem que

$$\frac{1}{3^n} = (0.\underbrace{0\dots 0}_{n-1}1)_3 = (0.\underbrace{0\dots 0}_n\widehat{2})_3.$$

Examinant com són aquest intervals es demostra que tenen la forma

$$I_k^n = [(0.b_1\dots b_n)_3, (0.b_1\dots b_n\widehat{2})_3],$$

amb $b_1, \dots, b_n \in \{0, 2\}$. Llavors, per (9),

$$\mathbf{P}(X \in I_k^n) = F((0.b_1\dots b_n\widehat{2})_3) - F((0.b_1\dots b_n)_3) = \sum_{j=n+1}^{\infty} \frac{2}{2^{j+1}} = \frac{1}{2^n}.$$

Comentaris

1. Pot demostrar-se que aquesta propietat caracteritza la distribució de Cantor. Concretament, si una variable aleatòria Y compleix que per qualsevol $n \geq 1$,

$$\mathbf{P}(Y \in I_k^n) = \frac{1}{2^n}, \quad k = 1, \dots, 2^n,$$

llavors Y té la distribució de Cantor.

2. La propietat (17), juntament amb altres que anirem veiem, justifiquen que de manera informal es digui que *la distribució de Cantor es comporta com una distribució uniforme sobre el conjunt de Cantor*. Una propietat de la distribució uniforme $[0, 1]$ és que si dividim l'interval en m subintervals de la mateixa longitud, $[0, 1/m)$, $[1/m, 2/m)$, \dots , $[(m-1)/m, 1]$, aleshores tots tenen la mateixa probabilitat.

8.3 La variable aleatòria $1 - X$ també té distribució de Cantor

Si X té una distribució de Cantor, aleshores la variable $Y = 1 - X$ també. Per demostrar-ho, anomenem G la funció de distribució de Y i volem veure que $G = F$. En efecte (recordeu de nou que F és contínua),

$$\begin{aligned} G(x) &= \mathbf{P}(Y \leq x) = \mathbf{P}(1 - X \leq x) = \mathbf{P}(1 - x \leq X) \\ &= 1 - \mathbf{P}(X < 1 - x) = 1 - F(1 - x) = F(x), \end{aligned} \quad (18)$$

on hem aplicat la propietat 6.5 de F .

Observació. Aquesta propietat equival a que **la distribució de Cantor és simètrica respecte el punt $1/2$** . Això vol dir que si dos conjunts són simètrics respecte el punt $1/2$, llavors tenen la mateixa probabilitat. Per exemple, si tenim un interval $[a, b]$, amb $0 < a < b < 1/2$, l'interval simètric respecte $1/2$ és $[1 - b, 1 - a]$, vegeu la Figura 8. Llavors,

$$\mathbf{P}(X \in [a, b]) = \mathbf{P}(X \in [1 - b, 1 - a]).$$

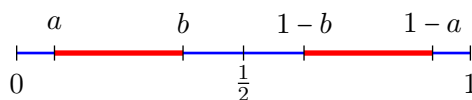


Figura 8: Els intervals $[a, b]$ i $[1 - b, 1 - a]$ són simètrics respecte el punt $1/2$

Per les propietats de les funcions de distribució, per demostrar que dos conjunts simètrics respecte $1/2$ tenen la mateixa probabilitat, n'hi ha prou amb comprovar que per a qualsevol $x \in [0, 1]$,

$$\mathbf{P}(X \in [0, x]) = \mathbf{P}(X \in [1 - x, 1]),$$

que equival a veure que X i $1 - X$ tenen distribució de Cantor.

8.4 Autosemblança de la distribució de Cantor

Aquesta propietat reposa en el caràcter fractal del conjunt C . Diu que si seleccionem a l'atzar un dels intervals I_1^1 i I_2^1 que formen C_1 , i hi prenem un punt d'acord amb la distribució de Cantor, tornem a obtenir la distribució de Cantor. En fórmules: siguin X_1 i X_2 dues variables aleatòries independents, ambdues amb distribució de Cantor, i definim la variable aleatòria Y per:

$$Y = \begin{cases} \frac{X_1}{3}, & \text{amb probabilitat } \frac{1}{2}, \\ \frac{X_2}{3} + \frac{2}{3}, & \text{amb probabilitat } \frac{1}{2}. \end{cases}$$

Aleshores Y té una distribució de Cantor. Noteu que $X_1/3 \in I_1^1$ i $X_2/3 + 2/3 \in I_2^1$. Alguns autors consideren aquesta propietat tan intuïtiva que no cal fer-ne cap demostració.

Per escriure de manera més compacta l'expressió anterior, introduïm una variable aleatòria que representi l'elecció a l'atzar entre I_1^1 i I_2^1 . Concretament, sigui Z una variable tal que

$$\mathbf{P}(Z = 0) = \mathbf{P}(Z = 1) = \frac{1}{2},$$

independent de X_1 i X_2 ; quan $Z = 1$, elegim I_1^1 , i quan $Z = 0$ elegim I_2^1 :

$$Y = \begin{cases} \frac{X_1}{3}, & \text{si } Z = 1, \\ \frac{X_2}{3} + \frac{2}{3}, & \text{si } Z = 0. \end{cases}$$

O, escrit en una línia,

$$Y = Z \frac{X_1}{3} + (1 - Z) \left(\frac{X_2}{3} + \frac{2}{3} \right). \quad (19)$$

La propietat es formula de la següent manera:

Autosemblança de la distribució de Cantor. Amb les notacions anteriors, Y té una distribució de Cantor.

Comentaris

1. El nom *autosemblança* prové d'una propietat important del conjunt de Cantor, vegeu Haro [15, p. 166] i aquí en fem una versió probabilística. No s'ha de confondre amb la propietat d'*autosimilitud* de certs processos estocàstics. Vegeu Delgado et al. [7, pp. 133–134]

2. Tal com passa amb el conjunt de Cantor (Haro [15, p. 166]), a l'apartat 8.11 veurem que aquesta propietat caracteritza la distribució de Cantor.

3. Hi ha una propietat anàloga en cada nivell; per exemple, al segon nivell, amb probabilitat $1/4$ prenem una de les components de C_2 : $I_1^2, I_2^2, I_3^2, I_4^2$, i llavors un punt amb distribució de Cantor, aleshores tornem a tenir una distribució de Cantor.

Demostració de la propietat. Designem la funció de distribució de Y per G . Volem veure que $G = F$. Fixem $t \in [0, 1]$. Pel teorema de les probabilitats totals,

$$\begin{aligned} G(t) &= \mathbf{P}(Y \leq t) = \mathbf{P}(Z = 1) \mathbf{P}(Y \leq t | Z = 1) + \mathbf{P}(Z = 0) \mathbf{P}(Y \leq t | Z = 0) \\ &= \frac{1}{2} \mathbf{P}\left(\frac{X_1}{3} \leq t\right) + \frac{1}{2} \mathbf{P}\left(\frac{X_2}{3} + \frac{2}{3} \leq t\right) \\ &= \frac{1}{2} \mathbf{P}(X_1 \leq 3t) + \frac{1}{2} \mathbf{P}(X_2 \leq 3t - 2) \\ &= \frac{1}{2} F(3t) + \frac{1}{2} F(3t - 2) = F(t), \end{aligned}$$

on per obtenir l'última igualtat s'ha de distingir segons t estigui a $[0, 1/3]$, $(1/3, 2/3)$ o $[2/3, 1]$, aplicar les propietats (12) de la funció F i que $F(t) = 1/2$ quan $t \in (1/3, 2/3)$.

8.5 Moments d'una variable aleatòria

Sigui Y és una variable aleatòria i considerem un nombre natural $n \geq 1$. Recordem que s'anomena el **moment d'ordre n** de Y a l'esperança de Y^n : $\mathbf{E}[Y^n]$, on es suposa que $\mathbf{E}[|Y|^n] < \infty$; si $\mathbf{E}[|Y|^n] = \infty$, es diu que Y no té moment d'ordre n . El moment d'ordre 1 s'anomena l'esperança de Y .

Els casos més habituals són quan la variable aleatòria Y és discreta o quan té densitat, i llavors es poden calcular els moments per fórmules (la qual cosa no vol dir que sempre sigui fàcil). Concretament, si Y és discreta, és a dir, hi ha una conjunt finit o numerable $B = \{y_j, j \in J\}$, amb $J \subset \mathbb{N}$, tal que $\mathbf{P}(Y \in B) = 1$, llavors

$$\mathbf{E}[Y^n] = \sum_{j \in J} y_j^n \mathbf{P}(Y = y_j),$$

sempre que $\sum_{j \in J} |y_j|^n \mathbf{P}(Y = y_j) < \infty$.

Quan Y té funció de densitat f , aleshores

$$\mathbf{E}[Y^n] = \int_{-\infty}^{\infty} y^n f(y) dy,$$

sempre que $\int_{-\infty}^{\infty} |y|^n f(y) dy < \infty$.

8.6 Moments de la distribució de Cantor

Sigui X una variable aleatòria amb la distribució de Cantor. Atès que

$$0 \leq X \leq 1,$$

per a qualsevol $n \geq 1$,

$$\mathbf{E}[|X|^n] \leq 1 < \infty.$$

Per tant, X tindrà moment d'ordre n .

Però hem vist a (15) que per a qualsevol $x \in \mathbb{R}$,

$$\mathbf{P}(X = x) = 0,$$

i llavors X no és discreta. (Seria més correcte dir que X no té part de salts).

D'altra banda, també hem comentat (veieu la propietat 4 de la secció 6) que per a tot $x \in \mathbb{R} - C$, $F'(x) = 0$, la qual cosa implica que X no té funció de densitat, és a dir, no hi ha cap funció f tal que per a qualsevol $a < b$,

$$P(a \leq X \leq b) = \int_a^b f(t) dt.$$

En efecte, si existís una funció f que complís la propietat anterior, òbviament no podria ser sempre zero (quasi per tot arreu), ja que llavors la probabilitat de l'esquerra seria sempre 0. Però també tindriem

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(t) dt,$$

i llavors $F'(x) = f(x) \neq 0$ (quasi per tot arreu), que és contradictori. Tècnicament es diu que X no té part absolutament contínua.

Aleshores, per calcular els moments de X cal buscar altres maneres de fer-ho.

8.7 Càlcul de l'esperança de la distribució de Cantor

Utilitzarem que (vegeu 8.3) X i $1 - X$ tenen ambdues distribucions de Cantor i per tant tindran la mateixa esperança:

$$\mathbf{E}[X] = \mathbf{E}[1 - X],$$

d'on s'obté que

$$\mathbf{E}[X] = 1/2.$$

8.8 Càlcul del moment de 2n. ordre i de la variància de la distribució de Cantor

Aquí utilitzarem la propietat d'autosemblança de l'apartat 8.4. Amb les notacions d'aquell apartat, X i $ZX_1/3+(1-Z)(X_2/3+2/3)$ tenen els mateixos moments. Utilitzant la independència entre totes les variables, que $Z(1-Z) = 0$, que X , X_1 i X_2 tenen la mateixa distribució, i que Z i $1-Z$ tenen la mateixa distribució, tenim,

$$\begin{aligned}\mathbf{E}[X^2] &= \mathbf{E}\left[\left(Z\frac{X_1}{3} + (1-Z)\left(\frac{X_2}{3} + \frac{2}{3}\right)\right)^2\right] \\ &= \frac{1}{9}\mathbf{E}[Z^2]E[X^2] + \frac{1}{9}\mathbf{E}[Z^2]\mathbf{E}[(X+2)^2] \\ &= \frac{1}{18}\mathbf{E}[X^2] + \frac{1}{18}\left(\mathbf{E}[X^2] + 4\mathbf{E}[X] + 4\right) \\ &= \frac{1}{18}\mathbf{E}[X^2] + \frac{1}{18}\mathbf{E}[X^2] + \frac{1}{3},\end{aligned}$$

d'on aïllant $\mathbf{E}[X^2]$ s'obté

$$\mathbf{E}[X^2] = \frac{3}{8}.$$

A partir d'aquí podem calcular la variància de X :

$$\text{Var}(X) = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 = \frac{3}{8} - \frac{1}{4} = \frac{1}{8}.$$

8.9 Moments d'ordre superior. Una recurrència pels moments de la distribució de Cantor

Amb el mateix mètode que hem calculat el moment de 2n ordre de la distribució de Cantor podem demostrar una fórmula de recurrència que ens permet calcular el moment de qualsevol ordre. Designem per m_n el moment d'ordre n de la distribució de Cantor:

$$m_n = \mathbf{E}[X^n], \quad n \geq 1.$$

Anem a provar la següent fórmula de recurrència (Lad and Taylor [16]):

$$m_{n+1} = \frac{1}{3^{n+1}-1} \sum_{j=0}^n \binom{n+1}{j} 2^{n-j} m_j, \quad n \geq 0. \quad (20)$$

La demostració és anàloga a la que hem fet del moment de 2n. ordre:

$$\begin{aligned}m_{n+1} &= \mathbf{E}\left[\left(Z\frac{X_1}{3} + (1-Z)\left(\frac{X_2}{3} + \frac{2}{3}\right)\right)^{n+1}\right] \\ &= \frac{1}{3^{n+1}}\mathbf{E}[Z^{n+1}X_1^{n+1} + (1-Z)^{n+1}(X_2+2)^{n+1}] \\ &= \frac{1}{2 \cdot 3^{n+1}}\left(m_{n+1} + \sum_{j=0}^{n+1} \binom{n+1}{j} m_j 2^{n+1-j}\right),\end{aligned}$$

i aïllant m_{n+1} s'obté (20).

Per exemple, a partir de $m_0 = 1$, tenim

$$m_1 = \frac{1}{2}, \quad m_2 = \frac{3}{8}, \quad m_3 = \frac{5}{16} \quad \text{i} \quad m_4 = \frac{87}{320}. \quad (21)$$

8.10 Funcions característiques

Entre les eines més importants de la Teoria de la probabilitat sobresurten les funcions característiques. Resultats molt importants, com les versions generals del Teorema central del límit, no van poder ser establerts fins que Paul Lévy va desenvolupar completament l'estudi d'aquestes funcions durant els anys 30 i 40 del segle XX.

La funció característica d'una variable aleatòria Y és la funció de variable real a valors complexos $\varphi : \mathbb{R} \rightarrow \mathbb{C}$ definida per

$$\varphi(t) = \mathbf{E}[e^{itX}], \quad t \in \mathbb{R}.$$

Tal com el seu nom indica, la funció característica caracteritza una distribució de probabilitat: hi ha una bijecció entre les funcions característiques i les distribucions de probabilitat, i moltes propietats de les distribucions poden estudiar-se a través de les corresponents funcions característiques.

Amb les notacions de l'apartat 8.5, quan Y és discreta,

$$\varphi(t) = \sum_{j \in J} e^{iy_j t} \mathbf{P}(Y = y_j).$$

Quan Y té funció de densitat f , aleshores

$$\varphi(t) = \int_{-\infty}^{\infty} e^{iyt} f(y) dy.$$

8.11 La funció característica de la distribució de Cantor

Per calcular la funció característica de la distribució de Cantor tornarem a utilitzar la propietat de l'apartat 8.4. Recordem que, d'una banda, X és variable aleatòria amb distribució de Cantor; de l'altra banda, X_1 i X_2 són dues variables aleatòries independents, ambdues amb distribució de Cantor, i Z una variable que pren els valors 0 i 1 amb probabilitat 1/2, independent de X_1 i X_2 . Aleshores $Z X_1/3 + (1-Z)(X_2/3 + 2/3)$ també té una distribució de Cantor.

Per la versió del Teorema de les probabilitats totals per a l'esperança,

$$\begin{aligned}
 \varphi(t) &= \mathbf{E}[e^{itX}] \\
 &= \mathbf{P}(Z = 1) \mathbf{E}\left[\exp\left\{it\left(Z X_1/3 + (1 - Z)(X_2/3 + 2/3)\right)\right\} \mid Z = 1\right] \\
 &\quad + \mathbf{P}(Z = 0) \mathbf{E}\left[\exp\left\{it\left(Z X_1/3 + (1 - Z)(X_2/3 + 2/3)\right)\right\} \mid Z = 0\right] \\
 &= \frac{1}{2} \mathbf{E}[e^{itX_1/3}] + \frac{1}{2} \mathbf{E}[e^{it(X_2/3+2/3)}] = \frac{1}{2}\varphi(t/3) + \frac{1}{2}e^{2it/3}\varphi(t/3) \\
 &= \frac{1}{2}\varphi(t/3)(1 + e^{2it/3}).
 \end{aligned}$$

A partir de la fórmula d'Euler

$$e^{ia} = \cos a + i \sin a,$$

tenim que

$$\cos a = \frac{1}{2}(e^{-ia} + e^{ia}) = \frac{1}{2}e^{-ia}(1 + e^{2ia}). \quad (22)$$

Aplicant-ho a $a = t/3$, tenim que

$$\varphi(t) = \varphi(t/3) e^{it/3} \cos(t/3).$$

Iterant aquest procediment arribem a

$$\varphi(t) = \varphi(t/3^n) \exp\left\{it \sum_{j=1}^n \frac{1}{3^j}\right\} \prod_{j=1}^n \cos(t/3^j).$$

Passant al límit quan $n \rightarrow \infty$ i utilitzant les propietats de les funcions característiques deduïm que

$$\varphi(t) = e^{it/2} \prod_{n=1}^{\infty} \cos(t/3^n). \quad (23)$$

Comentari

Hem esmentat a l'apartat 8.4 que la propietat d'auto semblança caracteritza la distribució de Cantor. Això vol dir el següent: sigui Y una variable aleatòria tal que si Y_1 i Y_2 són dues variables aleatòries independents, ambdues amb la mateixa distribució que Y , i Z una variable que pren els valors 0 i 1 amb probabilitat 1/2, independent de Y_1 i Y_2 , aleshores $Z Y_1/3 + (1 - Z)(Y_2/3 + 2/3)$ té la distribució de Y , llavors Y té la distribució de Cantor.

Aquesta propietat es dedueix del fet que només hem utilitzat la propietat d'auto semblança per calcular la funció característica de la distribució de Cantor. Llavors, qualsevol variable que compleixi la propietat d'auto semblança tindrà aquesta funció característica i, per tant, la distribució de Cantor.

9 La distribució de Cantor i les mesures de Hausdorff i de Haar

Per completesa comentarem breument la remarcable igualtat de la distribució de Cantor (restringida al conjunt C) amb les mesures de Hausdorff i de Haar. Aquesta secció és més avançada i suposarem que el lector coneix la terminologia estàndard de la teoria de la mesura tal com s'exposa, per exemple, a Schilling [25].

9.1 La distribució de Cantor a C

La funció de Cantor F , considerada com a funció de distribució, defineix directament una probabilitat \mathbf{P} sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, on $\mathcal{B}(\mathbb{R})$ és la σ -àlgebra de Borel sobre \mathbb{R} . Podem considerar la restricció de \mathbf{P} a C amb la σ -àlgebra de Borel sobre C , $\mathcal{B}(C)$, que coincideix amb la σ -àlgebra traça

$$\mathcal{B}(C) = \mathcal{B}(\mathbb{R}) \cap C = \{B \cap C, B \in \mathcal{B}(\mathbb{R})\}.$$

Designarem aquesta restricció per \mathbf{P}_C : per $A \in \mathcal{B}(C)$, de la forma $A = B \cap C$ per a algun $B \in \mathcal{B}(\mathbb{R})$ definim

$$\mathbf{P}_C(A) = \mathbf{P}(B \cap C).$$

Hem comentat a l'apartat 8.2 que la distribució de Cantor quedava determinada pel fet que tots els intervals $I_1^n, \dots, I_{2^n}^n$ tenien la mateixa probabilitat. Aquesta propietat es transllada a la restricció a C :

Caracterització de \mathbf{P}_C . Sigui \mathbf{Q} una probabilitat a $(C, \mathcal{B}(C))$ tal que per qualsevol $n \geq 1$,

$$\mathbf{Q}(I_k^n \cap C) = \frac{1}{2^n}, \quad k = 1, \dots, 2^n. \quad (24)$$

Aleshores $\mathbf{Q} = \mathbf{P}_C$.

9.2 Distribució de Cantor i mesura de Hausdorff

Aquesta secció està basada en Mattila [19, pp. 110]; per a la definició de mesura i dimensió de Hausdorff, el seu càlcul i interpretació veieu Falconer [11].

Considerem la mesura de Hausdorff de dimensió $s = \log 2 / \log 3$ a \mathbb{R} , que denotarem per \mathcal{H}^s , que és una mesura en $\mathcal{B}(\mathbb{R})$, però no una probabilitat, ja que $\mathcal{H}^s(\mathbb{R}) = \infty$. Però tenim (vegeu, per exemple, Falconer [11, p. 35])

$$\mathcal{H}^s(C) = 1. \quad (25)$$

Llavors, a l'igual com hem fet abans, podem considerar la restricció de \mathcal{H}^s a C , que designarem per $\mathcal{H}^s|_C$, i que gràcies a (25) sí que és una probabilitat a $\mathcal{B}(C)$. Tenim que

$$\mathbf{P}_C = \mathcal{H}^s|_C. \quad (26)$$

La demostració consisteix en veure que $\mathcal{H}^s|_C$ compleix la propietat d'equiprobabilitat en cada nivell que hem comentat a (24).

Assenyalem que Falconer [12, Theorem 6.6, p. 105]) diu que $\mathcal{H}^s|_C$ és «la mesura natural uniformement distribuïda a C », en concordança amb el que hem assenyalat sobre la uniformitat de la distribució de Cantor.

9.3 Distribució de Cantor i mesura de Haar

Per les definicions de grup topològic i mesura de Haar ens remetem a Lang [17, Capítol 16]. Per la mesura de Haar a C vegeu Naughton [20]. Un resultat fonamental en l'estudi dels grups topològics diu que, sota certes condicions, en un grup topològic existeix una mesura, anomenada **mesura de Haar**, única excepte una constant multiplicativa, que és invariant per l'esquerra per translacions.

En el cas concret de C , per definir-hi una llei de composició, comencem amb una operació a $\{0, 2\}$ que designarem per \oplus , donada per la taula

| | | |
|----------|---|---|
| \oplus | 0 | 2 |
| 0 | 0 | 2 |
| 2 | 2 | 0 |

Per a dos nombres $0.x_1x_2\dots$ i $0.y_1y_2\dots$ de C (no escric el subíndex 3) definim l'operació component a component:

$$0.x_1x_2\dots \oplus 0.y_1y_2\dots = 0.(x_1 \oplus y_1)(x_2 \oplus y_2)\dots$$

Es demostra que l'aplicació definida per aquesta operació és contínua, i òbviament també ho és l'aplicació determinada per el pas a l'invers ja que és la identitat. Llavors (es comproven les condicions tècniques) existeix la corresponent mesura de Haar.

Per veure que la distribució de Cantor sobre C coincideix amb la mesura de Haar (normalitzada a 1) cal provar que si denotem la translació segons $x \in C$ per

$$\mathbf{P}_x(A) = \mathbf{P}_C(x \oplus A), \quad A \in \mathcal{B}(C),$$

tenim que $\mathbf{P}_x = \mathbf{P}_C$. De nou, per demostrar-ho, es pot utilitzar la caracterització de la distribució de Cantor que hem vist a (24).

10 Simulació de la distribució de Cantor

La simulació d'una distribució de Cantor es fa pel mètode de la transformació inversa (vegeu, per exemple, Rubinstein and Kroese [23, p. 51]) que en aquest

cas consisteix en generar una variable aleatòria uniforme U en $(0, 1)$ i calcular $H^{-1}(U)$; en notació base 2, si $U = (0.b_1 \dots b_n)_2$ (vegeu (5)):

$$H^{-1}((0.b_1 \dots b_n)_2) = (0.a_1 \dots a_n)_3, \text{ on } a_j = 2b_j, j = 1, \dots, n.$$

Cal notar que el fet que H no sigui injectiva no suposa aquí cap problema ja que la no injectivitat era deguda als nombres amb dues expressions, una d'elles amb infinites xifres, i a la pràctica només s'obté un nombre finit de xifres.

A <https://stats.stackexchange.com/questions/229556/how-to-sample-from-cantor-distribution> es troba el següent programa en R que concreta aquesta idea de forma molt eficient:

```
binari.a.ternari <- function(x) {
  y <- 0
  x <- round(2^52 * x)
  for (i in 1:52) {
    y <- y + 2*(x %% 2) # a%%b és el restant de la divisió a/b
    y <- y/3
    x <- floor(x/2)    # floor és la part entera de x/2
  }
  y # y ja està en sistema decimal
}
n <- 1000
x <- runif(n)
y <- binari.a.ternari(x)
plot(ecdf(y), pch=".") # comprovació: funció de distribució empírica
```

11 Passejades aleatòries

En aquesta secció ens limitarem a comentar un parell d'exemples de passejades aleatòries. Per una introducció a les passejades aleatòries (i a la probabilitat) recomanem l'excel·lent llibre de Chung [4].

Exemple 1: una passejada bidimensional

Imaginem que passem per l'exemple de Barcelona.

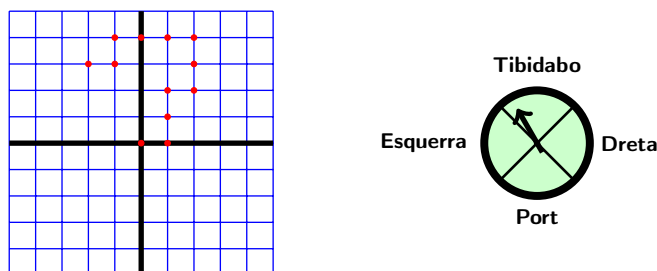


Figura 9: Una passejada aleatòria per l'exemple de Barcelona

Sortim d'un punt que situarem a l'origen i a cada cruïlla elegim aleatòriament la direcció, per exemple amb una ruleta com la del dibuix de la figura 9.

Exemple 2: una passejada unidimensional

Suposem ara que sortim de casa i a cada passa elegim aleatòriament si anem amunt o avall tirant una moneda. Vegeu la figura 10. A l'exemple, han sortit tres cares, una creu i dues cares

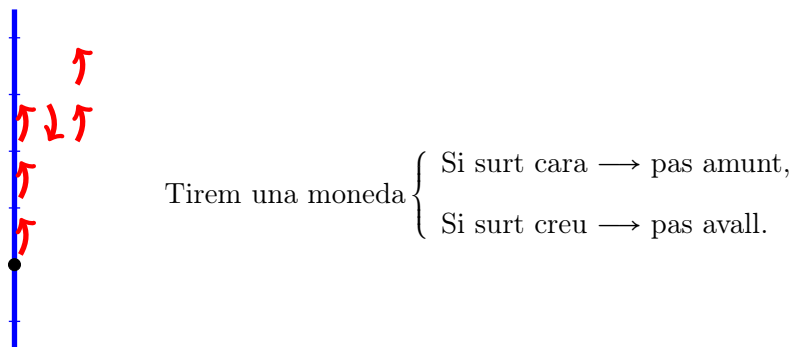


Figura 10: Una passejada unidimensional

Podem fer un gràfic espai-temps d'aquesta passejada aleatòria, vegeu la figura 11.

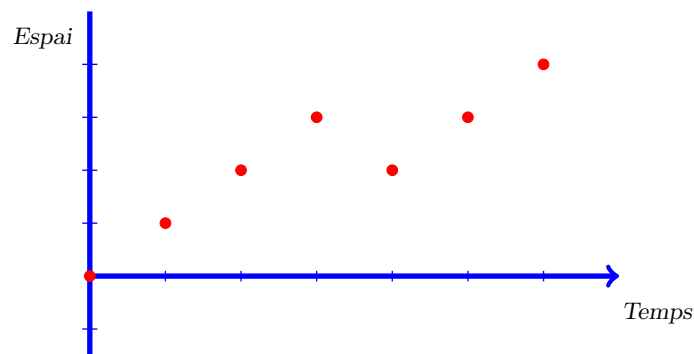


Figura 11: Gràfic espai-temps de la passejada unidimensional

Un model matemàtic de la passejada aleatòria unidimensional

Siguin Y_1, Y_2, \dots una successió de variables aleatòries independents, totes amb la mateixa distribució:

$$Y_n = \begin{cases} 1, & \text{amb probabilitat } \frac{1}{2}, \\ -1, & \text{amb probabilitat } \frac{1}{2}. \end{cases}$$

Sortim del zero: $Y_0 = 0$. La posició a l'instant k és

$$Y_1 + \dots + Y_k = \sum_{n=1}^k Y_n.$$

12 La distribució de Cantor com a límit d'una passejada aleatòria

Considerem una successió de variables aleatòries X_1, X_2, \dots independents i totes amb la següent distribució:

$$\mathbf{P}(X_j = 0) = \mathbf{P}(X_j = 2) = \frac{1}{2}.$$

Considerem la passejada aleatòria donada per la suma

$$\sum_{n=1}^k \frac{X_n}{3^n}.$$

En aquesta passejada els passos són cada cop són més curts. Es diu que és una passejada aleatòria amb passos decreixents.

La sèrie $\sum_{n=1}^{\infty} X_n/3^n$ convergeix absolutament q.s., ja que per a tot $\omega \in \Omega$, és una sèrie de termes positius i

$$\sum_{n=1}^{\infty} \frac{X_n(\omega)}{3^n} \leq 2 \sum_{n=1}^{\infty} \frac{1}{3^n} = 1 < \infty.$$

Definim S com la suma d'aquesta sèrie.

$$S = \sum_{n=1}^{\infty} \frac{X_n}{3^n}. \quad (27)$$

S és una variable aleatòria i demostrarem que té una distribució de Cantor.

12.1 Demostració que S té la distribució de Cantor

Designem per S_k la suma parcial:

$$S_k = \sum_{n=1}^k \frac{X_n}{3^n}.$$

Veurem que les funcions característiques de S_1, S_2, \dots , que denotarem $\varphi_{S_1}, \varphi_{S_2}, \dots$, respectivament, convergeixen a la funció característica de la distribució de Cantor φ , donada a (23), en tots els punts.

Per a $n = 1, 2, \dots$, la funció característica de X_n és

$$\varphi_{X_n}(t) = \mathbf{E}[e^{itX_n}] = \frac{1}{2}e^0 + \frac{1}{2}e^{2it} = \frac{1}{2}(1 + e^{2it}).$$

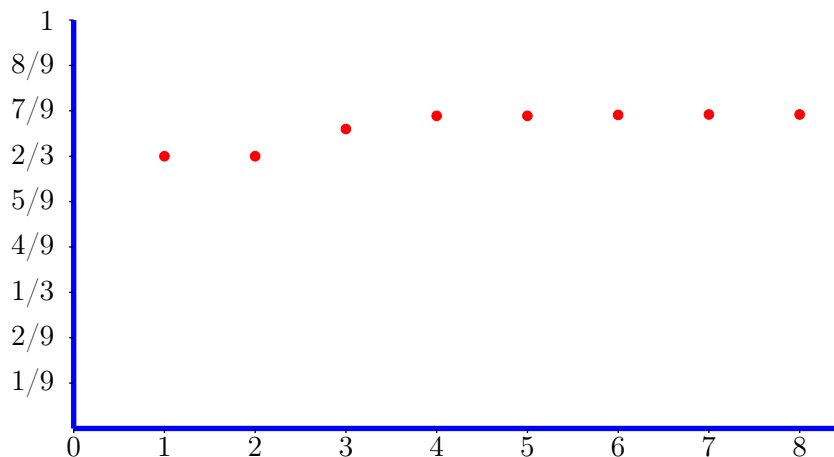


Figura 12: Una trajectòria de la passejada aleatòria S_1, S_2, \dots , que pren valors sobre el conjunt de Cantor

Llavors,

$$\varphi_{S_1}(t) = \varphi_{X_1/3}(t) = \varphi_{X_1}(t/3) = \frac{1}{2}(1 + e^{2it/3}).$$

La funció característica de $S_2 = X_1/3 + X_2/9$, gràcies a la independència entre X_1 i X_2 , serà

$$\varphi_{S_2}(t) = \varphi_{X_1}(t/3)\varphi_{X_2}(t/9) = \frac{1}{4}(1 + e^{2it/3})(1 + e^{2it/9}).$$

En general, la funció característica de S_n és

$$\varphi_{S_n}(t) = \frac{1}{2^n} \prod_{j=1}^n (1 + e^{2it/3^j}), \quad (28)$$

que gràcies a (22) es pot escriure

$$\varphi_{S_n}(t) = \exp \left\{ it \sum_{j=1}^n 3^{-j} \right\} \prod_{j=1}^n \cos(t/3^j).$$

Aleshores,

$$\lim_n \varphi_{S_n}(t) = e^{it/2} \prod_{n=1}^{\infty} \cos(t/3^n) = \varphi(t),$$

que coincideix amb (23). D'on resulta que la llei de S és la distribució de Cantor.

12.2 La distribució de Cantor com a límit de variables discretes uniformes

A partir de l'expressió (28), pot veure's que S_n té una llei uniforme en el conjunt dels extrems inferiors dels intervals I_k^n , $k = 1, \dots, 2^n$. Més concretament, si designem per $\mathcal{U}(c_1, \dots, c_k)$ una distribució uniforme sobre el conjunt

$\{c_1, \dots, c_k\}$, és a dir, cada punt amb probabilitat $1/k$, tindrem:

$$S_1 \sim \mathcal{U}(0, 2/3),$$

$$S_2 \sim \mathcal{U}(0, 2/3, 2/9, 8/9),$$

$$S_3 \sim \mathcal{U}(0, 2/27, 2/9, 8/27, 2/3, 20/27, 8/9, 26/27),$$

⋮

A la Figura 13 hi ha representada la funció de probabilitat de la variable S_3 i a la Figura 14 la seva funció de distribució. Hem vist que les funcions característiques de la successió S_1, S_2, \dots convergeixen a la funció característica de la distribució de Cantor; en conseqüència, les funcions de distribució corresponents convergiran a la funció de Cantor F . Compareu la Figura 14 amb la Figura 6 on hi ha el gràfic de la funció de distribució F_3 .

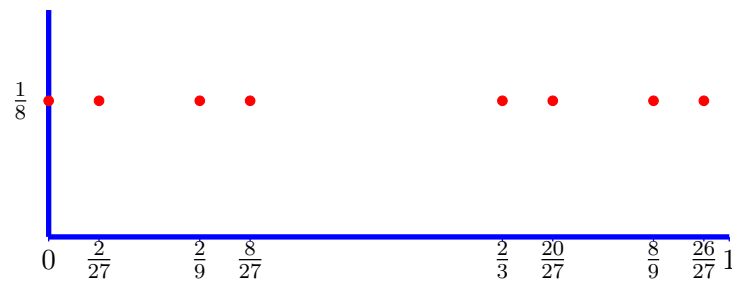


Figura 13: Funció de probabilitat de S_3

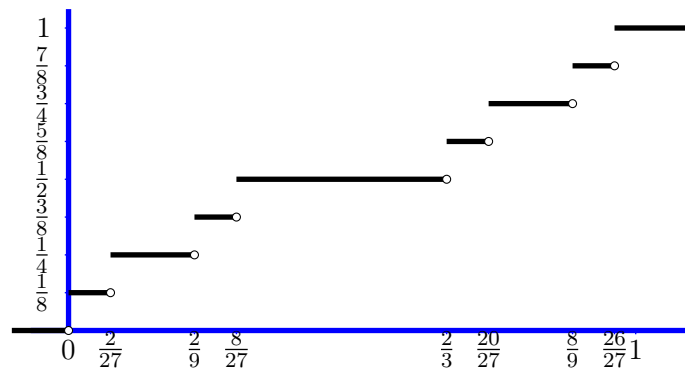


Figura 14: Funció de distribució de S_3

Així, la distribució de Cantor és el límit en distribució de variables discretes uniformes sobre certs punts de Cantor. Aquesta propietat ofereix una nova perspectiva a l'expressió que hem comentat abans sobre que la distribució de Cantor pot interpretar-se informalment com la distribució uniforme sobre el conjunt de Cantor.

13 Les propietats de la distribució de Cantor revisitades

Veiem que diverses propietats que hem estudiat de la distribució de Cantor són més intuïtives i fàcils de demostrar mitjançant la passejada aleatòria.

13.1 La distribució de Cantor està concentrada en el conjunt de Cantor

Per construcció, per a tot $\omega \in \Omega$, el valor $S(\omega)$ escrit en base 3 és una successió de 0 i 2, és a dir, són elements del conjunt de Cantor. Així, $\mathbf{P}(S \in C) = 1$.

13.2 Equiprobabilitat en cada nivell

Considerem un nivell arbitrari n . Hem comentat que els intervals I_k^n d'aquest nivell són de la forma

$$I_k^n = [(0.b_1 \dots b_n)_3, (0.b_1 \dots b_n \widehat{2})_3].$$

Aleshores, $S \in I_k^n$ equival a que $X_1 = b_1, \dots, X_n = b_n$. Llavors, per la independència entre X_1, \dots, X_n ,

$$\mathbf{P}(S \in I_k^n) = \mathbf{P}(X_1 = b_1, \dots, X_n = b_n) = \prod_{j=1}^n P(X_j = b_j) = \frac{1}{2^n}.$$

13.3 La distribució de Cantor és simètrica respecte el punt 1/2

Aquesta propietat es dedueix a partir del fet que les variables X_n i $2 - X_n$ tenen la mateixa distribució. Aleshores, les sèries

$$\sum_{n=1}^{\infty} \frac{X_n}{3^n} \quad \text{i} \quad \sum_{n=1}^{\infty} \frac{2 - X_n}{3^n}$$

tindran la mateixa distribució. Però la segona sèrie equival a

$$\sum_{n=1}^{\infty} \frac{2 - X_n}{3^n} = \sum_{n=1}^{\infty} \frac{2}{3^n} - \sum_{n=1}^{\infty} \frac{X_n}{3^n} = 1 - \sum_{n=1}^{\infty} \frac{X_n}{3^n},$$

és dir, que S i $1 - S$ tenen la mateixa distribució.

13.4 Moments de la distribució de Cantor

Els moments de la distribució de Cantor poden calcular-se a partir del fet la sèrie (27) és de termes positius, i aquestes sèries tenen molt bones propietats;

en particular, aplicarem la següent: si Y_n , $n \geq 1$, són variables positives, aleshores

$$\mathbf{E}\left[\sum_{n=1}^{\infty} Y_n\right] = \sum_{n=1}^{\infty} \mathbf{E}[Y_n].$$

Per a l'esperança, tenim que per a qualsevol n , $\mathbf{E}[X_n] = 0 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = 1$. Llavors

$$\mathbf{E}[S] = \sum_{n=1}^{\infty} \frac{\mathbf{E}[X_n]}{3^n} = \sum_{n=1}^{\infty} \frac{1}{3^n} = \frac{1}{2}.$$

Pel moment de 2n ordre partim de $\mathbf{E}[X_n^2] = 0 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{2} = 2$. Aleshores,

$$\mathbf{E}[S^2] = \sum_{n=1}^{\infty} \frac{\mathbf{E}[X_n^2]}{3^{2n}} + 2 \sum_{1 \leq n < m < \infty} \frac{\mathbf{E}[X_n X_m]}{3^{n+m}} = 2 \sum_{n=1}^{\infty} \frac{1}{3^{2n}} + 2 \sum_{1 \leq n < m < \infty} \frac{1}{3^{n+m}},$$

on hem utilitzat la independència entre X_n i X_m . La primera sèrie del terme de la dreta és la suma d'una progressió geomètrica de raó 1/9, i llavors,

$$\sum_{n=1}^{\infty} \frac{1}{3^{2n}} = \frac{\frac{1}{9}}{1 - \frac{1}{9}} = \frac{1}{8}.$$

La segona sèrie del terme de la dreta pot sumar-se utilitzant la fórmula (3) i, un altre cop, la suma d'una progressió geomètrica de raó 1/9:

$$\begin{aligned} \sum_{1 \leq n < m < \infty} \frac{1}{3^{n+m}} &= \sum_{n=1}^{\infty} \sum_{m=n+1}^{\infty} \frac{1}{3^{n+m}} \\ &= \sum_{n=1}^{\infty} \frac{1}{3^n} \sum_{m=n+1}^{\infty} \frac{1}{3^m} = \frac{1}{2} \sum_{n=1}^{\infty} \frac{1}{3^{2n}} = \frac{1}{2} \cdot \frac{1}{8} = \frac{1}{16}. \end{aligned}$$

Aleshores, tenim,

$$\mathbf{E}[S^2] = 2 \cdot \frac{1}{8} + 2 \cdot \frac{1}{16} = \frac{3}{8},$$

tal com havíem obtingut a l'apartat 8.8.

Aquest arguments poden generalitzar-se per calcular el moment de qualsevol ordre de S . Lad and Taylor [16] donen la fórmula següent: per a qualsevol nombre natural $N \geq 1$,

$$\mathbf{E}[S^N] = 2^N \sum_{k=1}^N \frac{1}{2^k} \sum \frac{N!}{a_1! \cdots a_k!} \frac{1}{\prod_{j=1}^k (3^{\sum_{i=j}^k a_i} - 1)},$$

on la segona suma es fa sobre totes les k -ples (a_1, \dots, a_k) de nombres naturals més grans o iguals a 1, tals que $N = \sum_{j=1}^k a_j$. Vegeu la Taula 1 amb els casos $N = 1, 2, 3, 4$.

| N | k | (a_1, \dots, a_k) |
|-----|-----|--------------------------|
| 1 | 1 | (1) |
| 2 | 1 | (2) |
| | 2 | (1,1) |
| 3 | 1 | (3) |
| | 2 | (1,2), (2,1) |
| | 3 | (1,1,1) |
| 4 | 1 | (4) |
| | 2 | (3,1), (1,3), (2,2), |
| | 3 | (2,1,1) (1,2,1), (1,1,2) |
| | 4 | (1,1,1,1) |

Taula 1: k -ples (a_1, \dots, a_k) de nombres naturals, $a_1 \geq 1, \dots, a_k \geq 1$, tals que $N = \sum_{j=1}^k a_j$, per a $N = 1, 2, 3, 4$.

Referències

- [1] AGUADÉ, J., Apunts d'un curs de topologia elemental. Accessible per internet a <http://mat.uab.es/~aguade/ACTE-Aguade-ver1.2.pdf>, consultat el 27/03/2020.
- [2] BARTLE, R. G., Introducción al análisis matemático. Limusa, Mexico, 1980.
- [3] CANTOR, G., De la puissance des ensembles parfaits de points, Acta Math. 4 (1884) 381–392.
- [4] CHUNG, K. L., Teoria elemental de la probabilidad y de los procesos estocásticos. Editorial Reverté, Barcelona, 1983.
- [5] CHUNG, K. L., A Course in Probability Theory, 3th edition. Academic Press, San Diego, 2001.
- [6] COHEN, P. J., Set Theory and the Continuum Hypothesis. Dover Publications, Inc., Mineola, New York, 2008.
- [7] DELGADO, R., JOLIS, M. I UTZET, F., Mandelbrot i l'atzar, Butl. Soc. Catalana Mat. 27, 2 (2012) 121–160.
- [8] DOVGOSHEY, O., MARTIO, O., RYAZANOV, V. AND VUORINEN, M., The Cantor function. Expo. Math. 24 (2006) 1–37.
- [9] ERDŐS, P., On a family of symmetric Bernoulli convolutions. Amer. J. Math. 61 (1939) 974–975.

- [10] ERDŐS, P., On the smoothness properties of Bernoulli convolutions. *Amer. J. Math.* 62 (1940) 180–186.
- [11] FALCONER, K., *Fractal Geometry*, 2nd edition. John Wiley and Sons, Chichester, 2003.
- [12] FALCONER, K., *Techniques in Fractal Geometry*. John Wiley and Sons, Chichester, 1997.
- [13] FLERON, J. F., A note on the history of the Cantor Set and Cantor Function. *Math. Mag.* 67, 2 (1994) 136–140 .
- [14] HALMOS, P. R., *Measure Theory*. Springer Verlag, Berlin, 1974
- [15] HARO, A., Fractalitat, determinisme i caos en el conjunt de Cantor. *Butl. Soc. Catalana Mat.* 27, 2 (2012) 161–175.
- [16] LAD, F.R. AND TAYLOR, F.C.R., The moments of the Cantor distribution. *Statist. Probab. Lett.* 13, 4 (1992) 307–310.
- [17] LANG, S., *Real Analysis*. Addison-Wesley Publishing Co., London 1983.
- [18] MANDELBROT, B., *Los objetos fractales: forma, azar y dimensión*. Tusquets, Barcelona, 1989.
- [19] MATTILA, P., *Fourier Analysis and Hausdorff Dimension*. Cambridge University Press, Cambridge, 2015.
- [20] NAUGHTON, G. P., *Haar Measure on the Cantor Ternary Set*, Master Thesis, University of North Texas. Accessible per Internet, https://digital.library.unt.edu/ark:/67531/metadc504018/m2/1/high_res_d/1002779028-Naughton.pdf, consultat el 30/03/2020.
- [21] PERES, Y., SCHLAG, W. AND SOLOMYAK, B., Sixty years of Bernoulli convolutions (1999). Accessible per internet, <https://u.math.biu.ac.il/~solomyb/RESEARCH/sixty.pdf>, consultat el 12/11/2020.
- [22] PESIN, Y. AND CLIMENHAGA, V., *Lectures on Fractal Geometry and Dynamical systems*. AMS, Providence, 2009.
- [23] RUBINSTEIN, R. I. AND KROESE, D. P., *Simulation and the Monte Carlo Method*, 2nd edition. John Wiley and Sons, Inc., Hoboken, New Jersey, 2008.
- [24] SANZ I SOLÉ, M., *Probabilitats*. Edicions de la Universitat de Barcelona, Barcelona, 1999.
- [25] SCHILLING, R., *Measures, Integrals and Martingales*. Cambridge University Press, Cambridge, 2005.



Departament de Matemàtiques
Universitat Autònoma de Barcelona
utzet@mat.uab.cat

Publicat el 18 de desembre de 2020